Star-Agents: Automatic Data Optimization with LLM Agents for Instruction Tuning

Hang Zhou[1,2], Yehui Tang[2], Haochen Qin[2], Yujie Yang[2], Renren Jin[1],Deyi Xiong[1*], Kai Han[2*], Yunhe Wang[2*]

[1] College of Intelligence and Computing, Tianjin University
[2] Noah's Ark Lab

NEURAL INFORMATION PROCESSING SYSTEMS

Background: Instruction Tuning Data

The quality of instruction tuning data plays a pivotal role.

- The instruction-following capability of LLMs are primarily acquired through instruction tuning .
- Expert-driven data generation assures the production of high-quality instructions, the enormous volume of data necessary for effective training renders this method economically untenable.
- The utilization of LLMs to automatically generate instructions, thereby mitigating the reliance on costly human annotation

Step1: Generating Diverse Data

Agent-Pair: Utilizing a spectrum of LLMs, each trained with discrepant setting, facilitates the generation of varied responses to given instructions.

The Star-Agents framework strategically pairs different LLMs to rewrite the instructions in the seed dataset and generate new responses to increase the diversity. With agent-pair (A^I_j, A^R_k), a new instruction data can be generated as follows:

$$f_{j,k}\left(I_i,R_i\right)=\left(A_j^I(I_i),A_k^R(R_i)\right),$$

$$D\left(S_i\right)=\left\{f_{j_1,k_1}\left(S_i\right),\cdots,f_{j_M,k_M}\left(S_i\right)\mid\left(j_m,k_m\right)\sim p_{jk},m=1,2,\cdots,M\right\},$$

Step2: Evaluating Tailored Data via a Dual-model Strategy

Dual-model Evaluation.

The intricate examples may surpass the capabilities of small models and be harmful for model performance, despite the advantages of using complex data for large models.

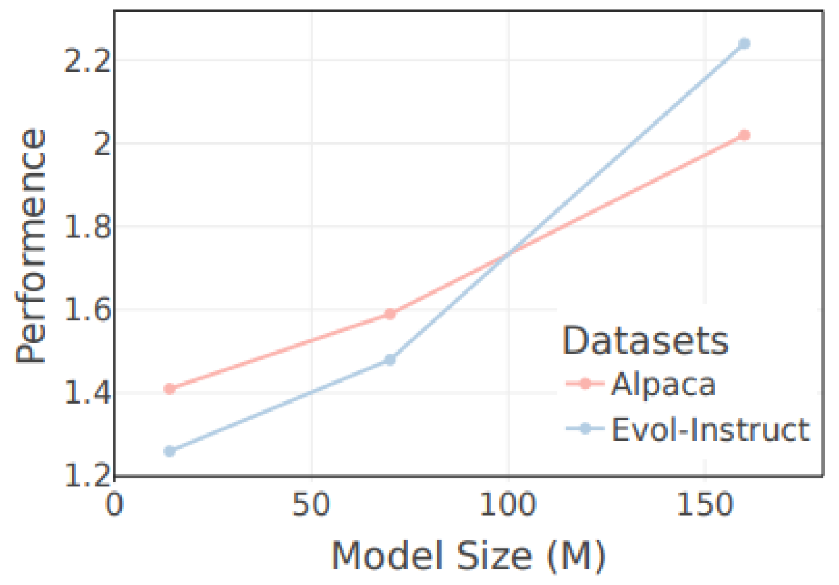


Figure 2: Performance comparison of varied scale models on the Alpaca and Evol-Instruc datasets. The tasks from the Evol-Instruct datase are more complex than those from Alpaca.

$$\mathrm{IFD}(I_i,R_i)=\frac{\exp\left(-\frac{1}{|R_i|}\sum_{w\in R_i}\log P(w|I_i)\right)}{\exp\left(-\frac{1}{|R_i|}\sum_{w\in R_i}\log P(w)\right)}.$$

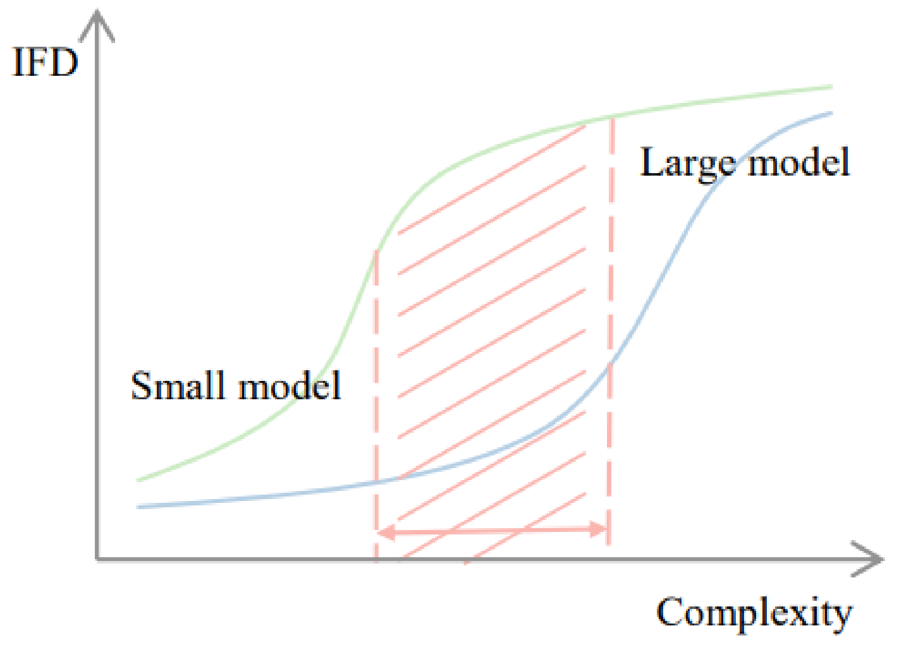
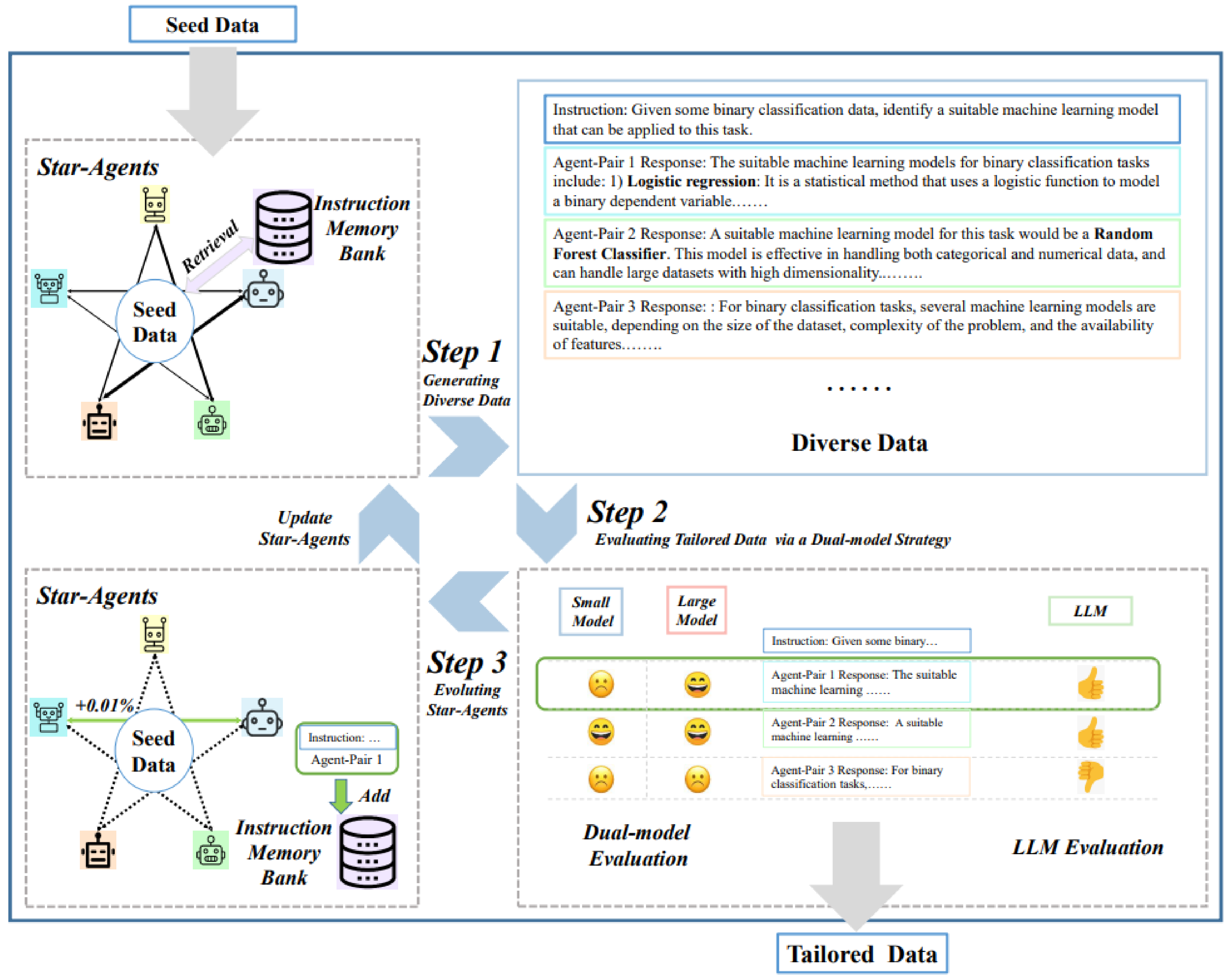


Figure 3: Illustration of dual-model evaluation. Data with a significant gap between the IFD scores of the small and large models will be prioritised.

We assume that for the same sample, stronger model yields a smaller IFD score. When the IFD scores of the two models are close to each other, it indicates that the sample is either too simple or too complex, which is not contributive to effective learning.

$$\pi_{\text{dual}}^{i}=\frac{\mathrm{IFD}_{\text{small}}(I_i,R_i)-\mathrm{IFD}_{\text{large}}(I_i,R_i)}{\max\limits_{1\le i\le m}\left(\mathrm{IFD}_{\text{small}}(I_i,R_i)-\mathrm{IFD}_{\text{large}}(I_i,R_i)\right)}.$$

Methods: Star-Agents


Seed Data

Star-Agents

Retrieval

Instruction Memory Bank

Seed Data

Step 1 Generating Diverse Data

Instruction: Given some binary classification data, identify a suitable machine learning model that can be applied to this task.

Agent-Pair 1 Response: The suitable machine learning models for binary classification tasks include: 1) **Logistic regression**: It is a statistical method that uses a logistic function to model a binary dependent variable.......

Agent-Pair 2 Response: A suitable machine learning model for this task would be a **Random Forest Classifier**. This model is effective in handling both categorical and numerical data, and can handle large datasets with high dimensionality.........

Agent-Pair 3 Response: : For binary classification tasks, several machine learning models are suitable, depending on the size of the dataset, complexity of the problem, and the availability of features........

......

Diverse Data

Update Star-Agents

Step 2 Evaluating Tailored Data via a Dual-model Strategy

Star-Agents

+0.01%

Seed Data

Instruction: ...
Agent-Pair 1

Add

Instruction Memory Bank

Step 3 Evoluting Star-Agents

Small Model

Large Model

LLM

Instruction: Given some binary...

Agent-Pair 1 Response: The suitable machine learning ......

Agent-Pair 2 Response: A suitable machine learning ......

Agent-Pair 3 Response: For binary classification tasks,......

Dual-model Evaluation

LLM Evaluation

Tailored Data

Figure 1: The diagram of the Star-Agents Framework. Step 1 is designed to gather diverse instructions and responses as shown in Appendix A.3. Step 2 focuses on selecting high-quality, tailored data from the data collected in Step 1. Finally, Step 3 aims to enhance the effectiveness and efficiency of the data generation process by evolving the Star-Agents framework.

Step3: Evolving Star Agents

Agent-Pair Sampling Evolution. The score π, which effectively estimates the quality of generated samples. During each iteration, if the generated samples are of high quality, we will increase the sampling probability of the selected agent-pair, which is updated as follows:

$$\tilde{p}_{jk}=p_{jk}+\beta\cdot\pi(I_i,R_i),$$

$$p_{jk}\leftarrow\frac{\tilde{p}_{jk}}{\sum_{j,k}\tilde{p}_{jk}}.$$

Instruction Memory Bank Evolution. We establish an Instruction Memory Bank storing highquality instructions aiming to accelerate sampling and relate the evolution with task data. When processing a data sample (I_i, R_i), we perform a query in the Instruction Memory Bank for I_i, retrieving the top n closest matches according to embedding similarity. The associated agent-pairs, identified as highly proficient for tasks similar to I_i , are then sampled. Subsequently, the Instruction Memory Bank will continuously evolve by incorporating tailored high-quality data

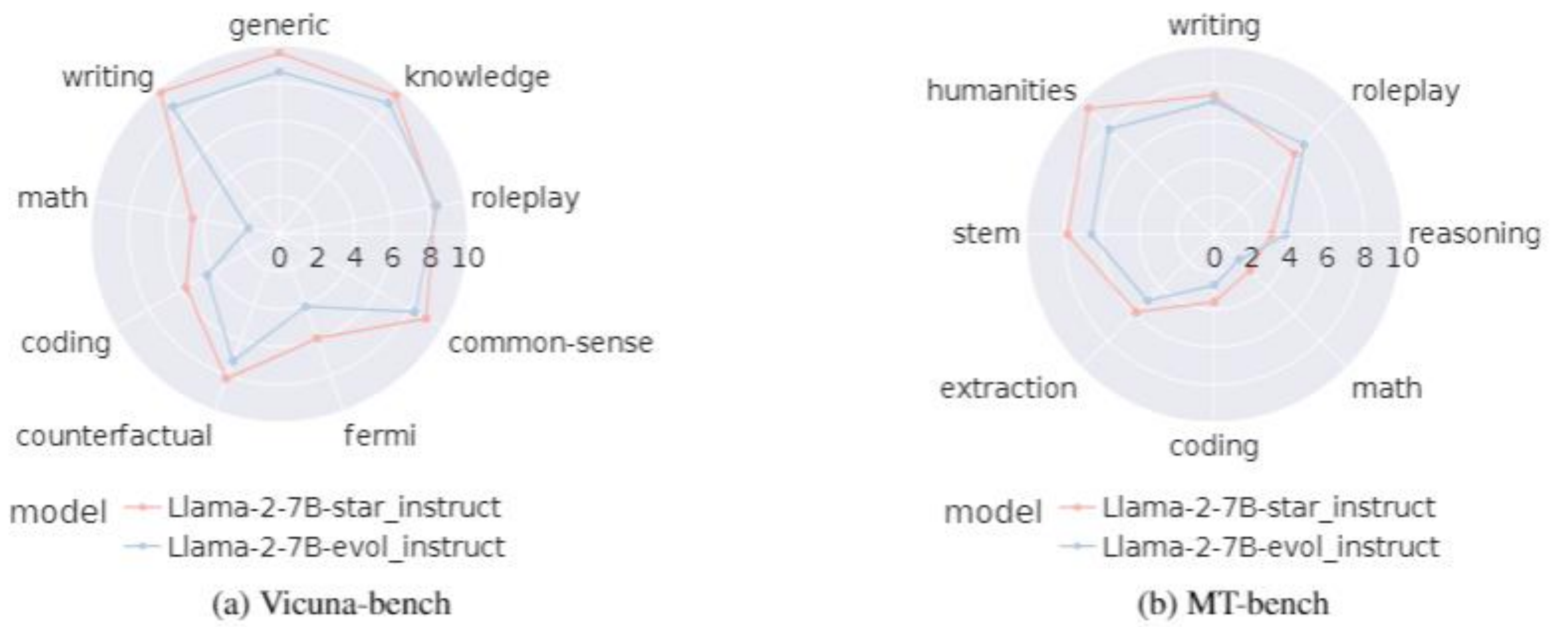
$$\pi_{\text{llm}}=\begin{cases}0, & \text{if the base data sample is better,}\\ 1, & \text{if the generated data sample is better,}\\ 0.5, & \text{if tie.}\end{cases}$$

$$\pi=\pi_{\text{llm}}\cdot\pi_{\text{dual}}.$$

Experiments

Table 2: Results of different models on Vicuna-bench, WizardLM testset and MT-Bench.

Model	Vicuna-Bench	WizardLM testset	MT-Bench	Average
1B Models				
Pythia-1B [2]	1.68	1.34	1.17	1.40
OPT-1.3B [47]	2.49	1.64	1.12	1.75
Sheared-LLaMA-1.3B [37]	2.73	1.86	1.59	2.06
Pythia-1B-alpaca	4.14	2.97	2.20	3.10
Pythia-1B-evol_instruct	5.07	3.55	2.56	3.73
Pythia-1B-IFD [19]	4.60	3.21	1.98	3.26
Pythia-1B-Random	5.13	3.39	2.35	3.62
Pythia-1B-star_instruct	**5.93**	**3.90**	**2.69**	**4.17**
7B Models				
Llama-2-7B [30]	-	-	3.95	-
zephyr-beta-sft [20]	-	-	5.32	-
mpt-7B-chat [20]	-	-	5.45	-
XGen-7B-8k-Inst [24]	-	-	5.55	-
sRecycled-Wiz-7B-v2 [16]	-	-	5.56	-
Llama-2-7B-alpaca	6.33	5.08	3.63	5.01
Llama-2-7B-evol_instruct	7.27	6.57	5.21	6.35
Llama-2-7B-star_instruct	**8.24**	**6.87**	**5.74**	**6.95**



(a) Vicuna-bench

(b) MT-bench

Figure 4: Radar plot of detailed scores for Llama-2-7B-star_instrcut against the major baseline on different subtasks of (a) Vicuna-Bench and (b) MT-Bench.

Table 3: Impact of different components.

Components			Average Score
Diversity	Data selection	Evolutiuon	
✓	✓	✓	**4.17**
✓	✓	✕	3.97
✓	✕	✕	3.62
✕	✕	✕	3.73

Table 4: Imapct of the Selection method.

Model	Average Score
Pythia-1B-evol_instruct	3.73
Pythia-1B-IFD [19]	3.26
Pythia-1B-Random	3.62
Pythia-1B-star_instruct	**4.17**